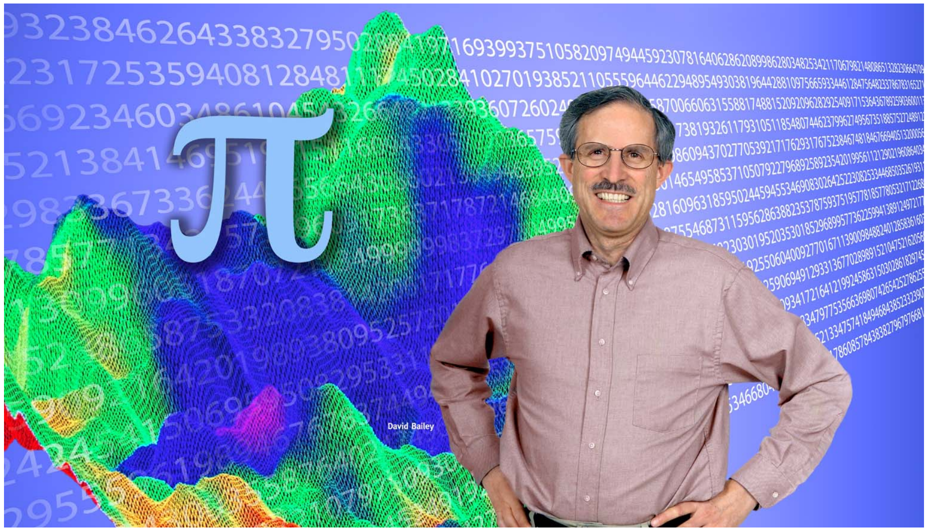
Normality and nonnormality of mathematical constants

David H Bailey

dhbailey@lbl.gov



Normal numbers

Given an integer $b > 1$, a real number α is ***b*-normal** (or "normal base b") if every m-long string of digits appears in the base-b expansion of α with precisely the expected limiting frequency $1 / b^m$.

Base-10 example: Suppose α satisfies all of these conditions:

1. Every digit appears, in the limit, with frequency 1/10.
2. Every 2-long string of digits (e.g., "23", "55", "74", etc.) appears, in the limit, with frequency 1/100.
3. Every 3-long string (e.g., "234", "551", "749", etc.) appears, in the limit, with frequency 1/1000.
4. Every 4-long string (e.g., "2345", "5518", "7493", etc.) appears, in the limit, with frequency 1/10000.

Similarly for **every** m-long string of digits, for **every** integer $m > 1$.

Then α is "10-normal" or "normal base 10."

Using measure theory, it is can be shown that, given an integer b > 1, almost all real numbers are b-normal. In fact, almost all reals are b-normal for all integer bases $b > 1$ simultaneously (i.e., are "absolutely normal").

Which specific reals are normal?

These are widely believed to be *b*-normal, for all integer bases $b > 1$:

π = 3.1415926535…

e = 2.7182818284…

sqrt(2) = 1.4142135623…

log(2) = 0.6931471805…

Every irrational algebraic number (this conjecture is due to Borel).

But there are **no proofs** of normality for any of these constants in any base, nor are there any nonnormality results.

Until recently, normality proofs were known only for a few relatively contrived examples, such as Champernowne's constant = 0.123456789101112131415… (which is 10-normal).

Chaitin's omega constant (from the theory of computational complexity) has been shown to be absolutely normal.

Fascination with π

Newton (1670):

- "I am ashamed to tell you to how many figures I carried these computations, having no other business at the time."

Carl Sagan (1986):

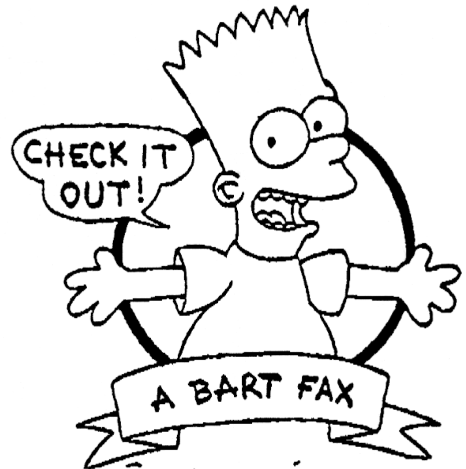
- In his book *Contact,* the lead scientist (played by Jodie Foster in the movie version) looked for patterns in the digits of π.

Carl E. Sagan

"Is π normal?" (and why) is one of the most ancient and fundamental of long-standing unsolved mathematical questions.

Gaining insight on this problem has been a prime motivation for many of the computations of π through the ages.

Fax to DHB from "The Simpsons" TV show



TO: DAVID BAILEY
FROM: JACQUELINE ATKINS
DATE: 10/9/92
NUMBER OF PAGES: 1

FAX (310) 203-3852
PHONE (310) 203-3959

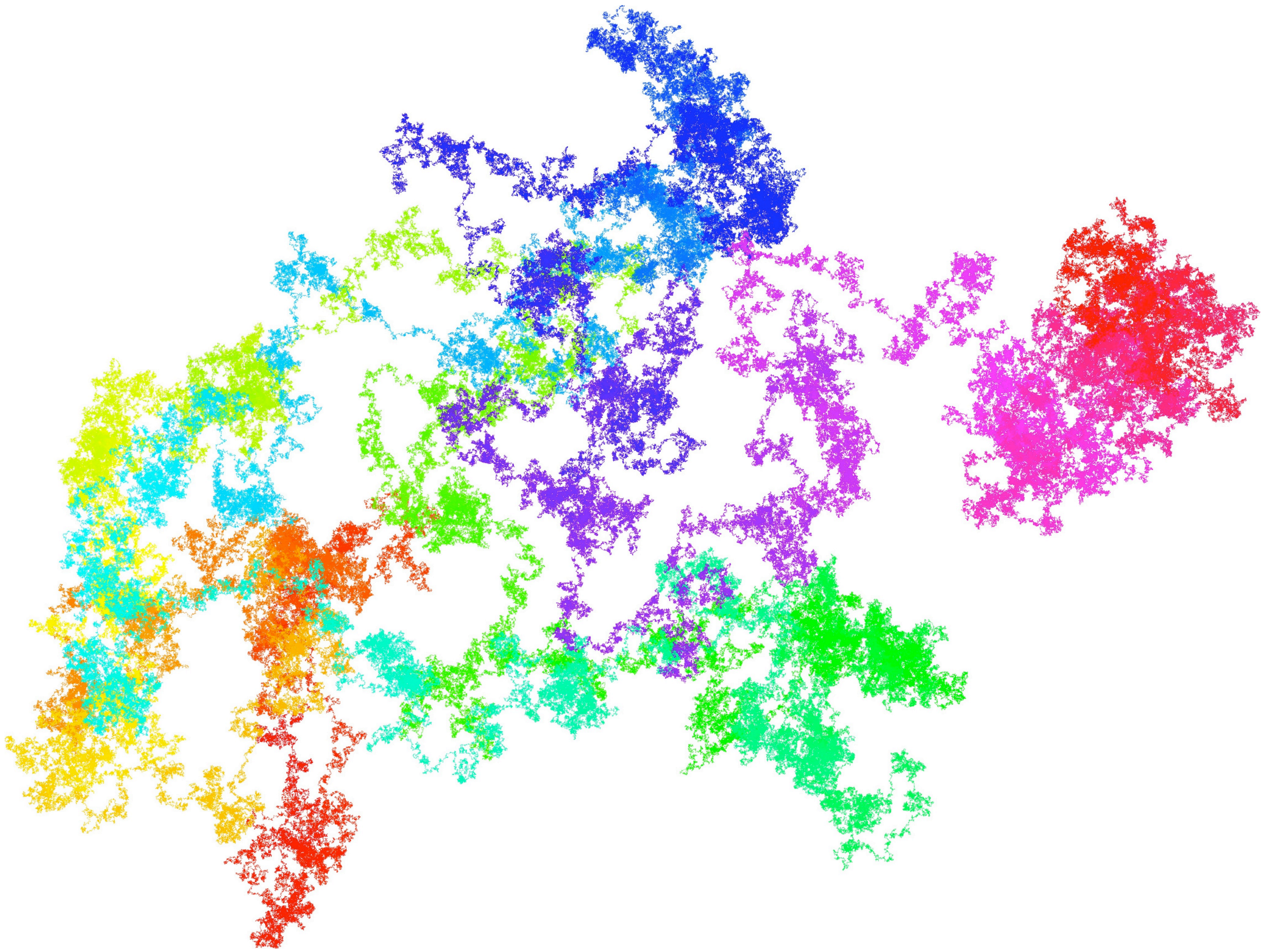
A Professor at UCLA told me that you might be able to give me the answer to: What is the 40,000th digit of Pi?

We would like to use the answer in our show. Can you help?

21st century techniques to address the normality problem

- Statistical analyses, using sophisticated models of normality:
  - D.H. Bailey, J.M. Borwein, C.S. Calude, M.J. Dinneen, M. Dumitrescu and A. Yee, “Normality and the digits of pi, http://crd-legacy.lbl.gov/~dhbailey/dhbpapers/normality-digits-pi.pdf.
- Graphics and visualization – e.g., viewing digits as a random walk:
  - F.J. Aragon Artacho, D.H. Bailey, J.M. Borwein and P.B. Borwein, “Walking on real numbers,” http://crd-legacy.lbl.gov/~dhbailey/dhbpapers/tools-walk.pdf.
- Analyses of algebraic irrationals:
  - D.H. Bailey, J.M. Borwein, R.E. Crandall and C. Pomerance, “On the binary expansions of algebraic numbers,” *Journal of Number Theory Bordeau*, vol. 16 (2004), pg. 487-518.
  - Hajime Kaneko, “On normal numbers and powers of algebraic numbers,” *Integers*, vol. 10 (2010), pg. 31–64.
- Connections to the theory of BBP-type constants:
  - D.H. Bailey and R.E. Crandall, "On the random character of fundamental constant expansions," *Experimental Mathematics*, vol. 10, no. 2 (Jun 2001), pg. 175-190.
  - D.H. Bailey and R.E. Crandall, “Random generators and normal numbers,” *Experimental Mathematics*, vol. 11, no. 4 (2002), pg. 527-546.
- Analyses of Stoneham constants:
  - D.H. Bailey and J.M. Borwein, “Nonnormality of Stoneham constants,” *Ramanujan Journal*, http://crd-legacy.lbl.gov/~dhbailey/dhbpapers/nonnormality.pdf.
  - D.H. Bailey and J.M. Borwein, “Normal numbers and pseudorandom generators,” http://crd.lbl.gov/~dhbailey/dhbpapers/normal-pseudo.pdf.

Random walk on the first two billion bits of π
(base-4 digit 0,1,2,3 codes up, down, left, right)



F.J. Aragon Artacho, D.H. Bailey, J.M. Borwein and P.B. Borwein, “Walking on real numbers,” available at http://crd-legacy.lbl.gov/~dhbailey/dhbpapers/tools-walk.pdf.

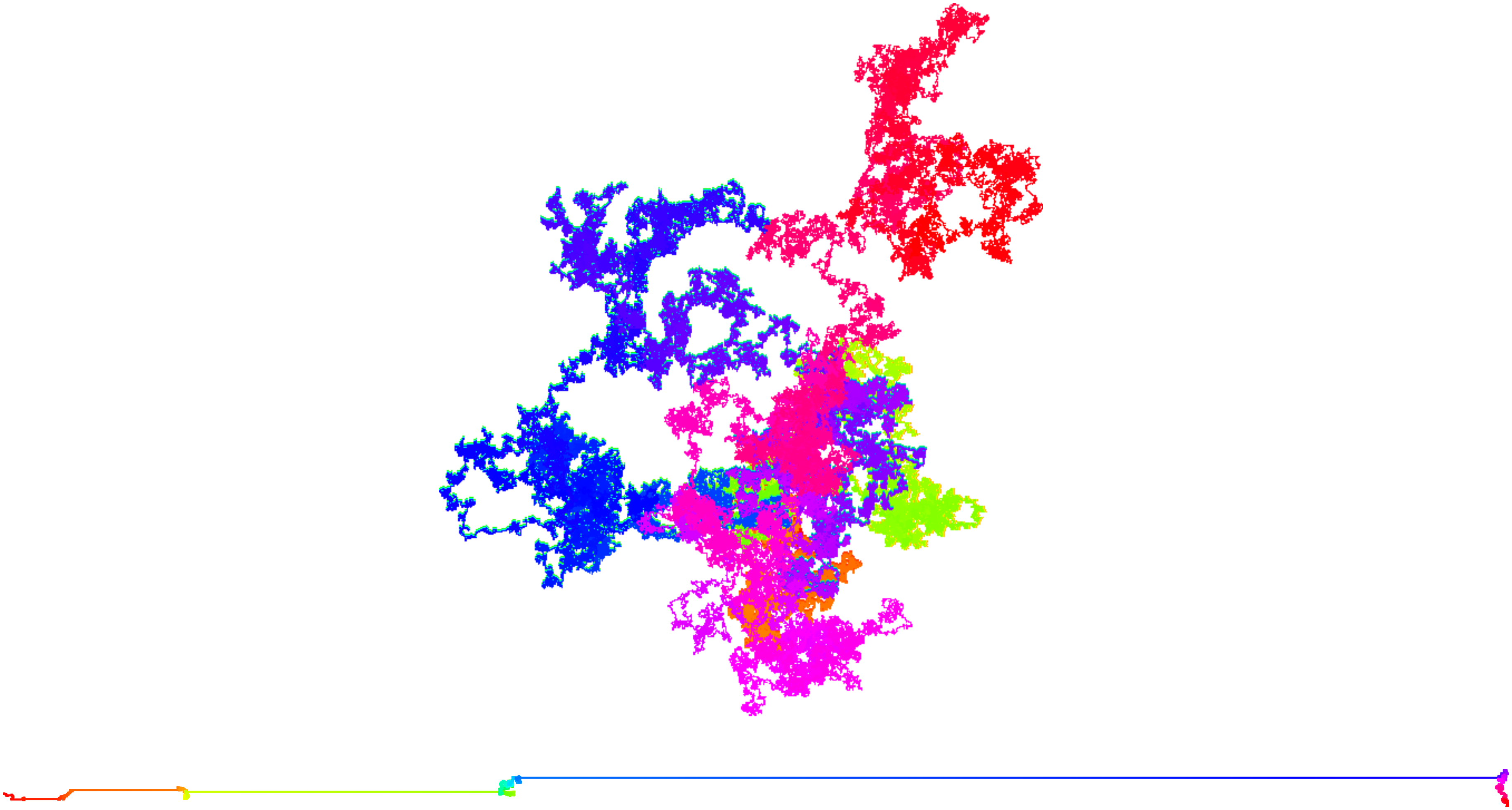
Explore the digits of π yourself

A compelling online interactive tool is available to study the random walk on the first 100 billion binary digits of π:

http://gigapan.com/gigapans?tags=pi

This was constructed by Fran Aragon Artacho, based on binary digits of π provided by Alex Yee.

Random walk on $\alpha_{2,3}$ base 2 (normal) compared to $\alpha_{2,3}$ base 6 (nonnormal)



We will discuss the normality and nonnormality of $\alpha_{2,3}$ later in the talk.

Poisson model analysis of digits of π

- Recently a Poisson model-based approach was employed to analyze the normality of π, using the first 15,925,868,541,400 bits (approx. 4 trillion hexadecimal digits) of π, as computed by Alex Yee.

- Based on this analysis, the conclusion "π is not normal" (given the first 4 trillion hexadecimal digits) has probability 10^{-3064}.

D.H. Bailey, J.M. Borwein, C.S. Calude, M.J. Dinneen, M. Dumitrescu and A. Yee, "Normality and the digits of pi," manuscript, http://crd-legacy.lbl.gov/~dhbailey/dhbpapers/normality-digits-pi.pdf.

A result for algebraic numbers

If x is algebraic of degree $d > 1$, then its binary expansion through position n must have at least $C\, n^{1/d}$ 1-bits, for all sufficiently large n and some C that depends on x.

Example: The first n binary digits of sqrt(2) must have at least sqrt(n) ones.

For the special case sqrt(m) for integer m, the result follows by simply noting that in binary notation, the one-bit count of the product of two integers is less than or equal to the product of the one-bit counts of the two integers.

A related result was obtained by Hajime Kaneko of Kyoto University.

However, note that these results are still a far cry from a full proof of normality, even in the single-digit binary sense.

D.H. Bailey, J.M. Borwein, R.E. Crandall and C. Pomerance, “On the binary expansions of algebraic numbers,” *Journal of Number Theory Bordeau*, vol. 16 (2004), pg. 487-518.

BBP formulas for π and log 2

In 1996, a computer program running the PSLQ algorithm discovered this formula, now known as the BBP formula for π:

$$\pi = \sum_{n=0}^{\infty} \frac{1}{16^n} \left(\frac{1}{8n+1} - \frac{1}{8n+4} - \frac{1}{8n+5} - \frac{1}{8n+6} \right)$$

This formula permits one to directly calculate binary or hexadecimal (base-16) digits of π beginning at an arbitrary starting position *n*, without needing to calculate any of the first *n*-1 digits.

A similar formula for log 2, which also has the arbitrary digit calculation property, is:

$$\log 2 = \sum_{n=1}^{\infty} \frac{1}{n 2^n}$$

D.H. Bailey, P.B. Borwein and S. Plouffe, “On the rapid computation of various polylogarithmic constants,” *Mathematics of Computation*, vol. 66, no. 218 (Apr 1997), pg. 903-913.

How to calculate individual binary digits of log 2

Let *x* mod 1 denote the fractional part of *x*. Note that we can write the binary expansion of log 2 beginning after position *d* as follows:

$$
\begin{aligned}
(2^d \log 2) \bmod 1 &= \left(\sum_{n=1}^{d} \frac{2^{d-n}}{n}\right) \bmod 1 + \left(\sum_{n=d+1}^{\infty} \frac{2^{d-n}}{n}\right) \\
&= \left(\sum_{n=1}^{d} \frac{2^{d-n} \bmod n}{n}\right) \bmod 1 + \left(\sum_{n=d+1}^{\infty} \frac{2^{d-n}}{n}\right)
\end{aligned}
$$

The numerator of first part can be computed very rapidly, using the binary algorithm for exponentiation (i.e., the observation that an exponentiation can be accelerated by using the binary expansion of the exponent). Only a few terms of the second part need be computed, since it is very small.

This “trick” also leads to an interesting connection to normality.

BBP formulas, pseudorandom number generators and normality

Consider the sequence $x_0 = 0$, and

$$x_n = \left(2x_{n-1} + \frac{1}{n}\right) \bmod 1$$

(The $1/n$ term comes from BBP formula for log 2). If it can be demonstrated that this sequence is equidistributed in the unit interval, then this would imply that log 2 is 2-normal.

Similarly, consider the sequence $x_0 = 0$, and

$$x_n = \left(16x_{n-1} + \frac{120n^2 - 89n + 16}{512n^4 - 1024n^3 + 712n^2 - 206n + 21}\right) \bmod 1$$

(The large term is a combination of the four terms in BBP formula for π). If it can be shown that this sequence is equidistributed in the unit interval, then this would imply that π is 16-normal (and hence 2-normal).

D.H. Bailey and R. E. Crandall, “On the random character of fundamental constant expansions,” *Experimental Mathematics*, vol. 10, no. 2 (Jun 2001), pg. 175-190.

Normality of the Stoneham constants

By extending the BBP approach, normality has been proved for the Stoneham numbers, the simplest of which is:

$$\begin{aligned}\alpha_{2,3} &= \sum_{n=1}^{\infty} \frac{1}{3^n 2^{3^n}} \\ &= 0.041883680831502985071252898624571682426096\ldots_{10} \\ &= 0.0ab8e38f684bda12f684bf35ba781948b0fcd6e9e0\ldots_{16}\end{aligned}$$

This particular constant was proven 2-normal by Stoneham in 1971. This has been extended to the case where (2,3) are any pair (*b*,*c*) of coprime integers > 1, and also to an uncountable class (here r_n is *n*-th bit of *r* in [0,1)):

$$\alpha_{2,3}(r) = \sum_{n=1}^{\infty} \frac{1}{3^n 2^{3^n + r_n}}$$

More recently, the 2-normality of $\alpha_{2,3}$ was proven more simply by means of a "hot spot" lemma proved using ergodic theory.

D.H. Bailey and M. Misiurewicz, "A strong hot spot theorem," *Proceedings of the American Mathematical Society*, vol. 134 (2006), no. 9, pg. 2495-2501.
D.H. Bailey and R.E. Crandall, "Random generators and normal numbers," *Experimental Mathematics*, vol. 11, no. 4 (2002), pg. 527-546.

The “hot spot” lemma and a proof of normality for $\alpha_{2,3}$

Let *A*(*a*,*y*,*n*,*m*) denote the count of occurrences where the *m*-long binary string *y* is found to start at position *p* in the base-*b* expansion of α, where 1 <= *p* <= *n*. Then if *x* is not *b*-normal, there is some *y* in the unit interval s.t.

$$\liminf_{m\to\infty} \limsup_{n\to\infty} \frac{b^m A(x,y,n,m)}{n} = \infty$$

Thus if α is not *b*-normal, there must exist some interval $[r_1, s_1)$ in which successive shifts of the base-b expansion of a visit $[r_1,s_1)$ 10 times more frequently, in the limit, relative to its length $s_1 - r_1$; there must be another interval $[r_2, s_2)$ that is visited 100 times more often relative to its length; etc.

On the other hand, if it can be established that no subinterval of the unit interval is visited, say, 1,000 times more often relative to its length by successive shifts of the base-*b* digits, this suffices to prove α is *b*-normal.

This can be condition is met for $\alpha_{2,3}$ in base-2. Thus $\alpha_{2,3}$ is 2-normal.

D.H. Bailey and M. Misiurewicz, “A strong hot spot theorem,” *Proceedings of the American Mathematical Society*, vol. 134 (2006), no. 9, pg. 2495-2501.

A pseudorandom number generator based on the binary digits of $\alpha_{2,3}$

- It can be shown that one can generate successive sections of the binary expansion of $\alpha_{2,3}$ by means of a simple scheme analogous to commonly used linear congruential pseudorandom number generators.
- The period of this generator is 3.7×10^{15}. Higher periods are possible if one uses higher-precision arithmetic.
- Floating-point numbers so generated appear to have excellent statistical properties, and do not have problems with power-of-two strides.
- The generator has a “jump-ahead” feature, which is useful in a parallel environment so each processor can generator its own section.
- The generator can be implemented at a speed matching that of other commonly used pseudorandom generators.

D.H. Bailey and J.M. Borwein, “Normal numbers and pseudorandom generators,”
http://crd.lbl.gov/~dhbailey/dhbpapers/normal-pseudo.pdf.

A nonnormality result

Although $\alpha_{2,3}$ is provably 2-normal, surprisingly it is NOT 6-normal. Note that we can write

$$6^n \alpha_{2,3} \bmod 1 = \left(\sum_{m=1}^{\lfloor \log_3 n \rfloor} 3^{n-m} 2^{n-3^m} \right) \bmod 1 + \sum_{m=\lfloor \log_3 n \rfloor + 1}^{\infty} 3^{n-m} 2^{n-3^m}$$

The first portion of this expression is zero, since all of the terms in the summation are integers. When $n = 3^m$, the second part is accurately approximated by the first term of the series. Thus,

$$6^{3^m} \alpha_{2,3} \bmod 1 \approx \frac{\left(\frac{3}{4}\right)^{3^m}}{3^{m+1}}$$

Since this is extremely small for large m, the base-6 expansion of $\alpha_{2,3}$ has long stretches of zeroes beginning at positions $3^m + 1$. This observation can be fashioned into a rigorous proof of nonnormality.

D.H. Bailey and J. M. Borwein, "Normal numbers and pseudorandom generators," available at http://crd.lbl.gov/~dhbailey/dhbpapers/normal-pseudo.pdf.
D.H. Bailey and J.M. Borwein, "Nonnormality of Stoneham constants," *Ramanujan Journal*, http://crd-legacy.lbl.gov/~dhbailey/dhbpapers/nonnormality.pdf.

Base-6 expansion of $\alpha_{2,3}$

0.
0130140430003334251130502130000001243555045432233011500243525320551352343541010
4300000000000000000514113005404055545530314425043343510124134523511251421251345
0550354501505352205204434045215150510241155250042513005112445400104413115003242
0303213000000000000000000000000000000000001421203431112145201352544534211341224
0220525301054204423552411055415015520435041455540031014530303353200253434040130
1240104453254343502142020432415025555101004043300045544114501031331451151014451
4123443342341240055131333504542353055315111535015334524354502500555214530542343
4215303501250242054041354512313232453530315345523041150201542421211452015422225
3434034045053012332553444044310333244533214141501423345454241243203125340050134
1502455144043000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000313350542444431111055534141052
0145402134123130014243331331155114400255244334414300322433530010235351531040203
1320020533130021221200323413302222122521000035254543002545040035245112205252324
3002003325231040225223410335201122534120341054402053234510110200235251500152125
4550242154534313455323500154554040002341152134232333535244525403554145221413251
4501442052123331155522222331511535031203434310450421303512422055300241415240113
0132422412341520333012053053004100454115523233154041243354304123023022441352550
1152231342154301320510050550243034442211503200331204300035524035115501414131024
4412135431134045402354431451021050554450430513450125125412054431444230051425001
1231544125122115413554321234142140304435152040242355245245543005025313140110150
4124231353020000520243340203202541130132402232021314323535151055043554353405331
4401324313010031034432330333020541411524554102024511252554525045245533541442124
0512042420440210023151214412155535320444505025514520102410432420550234114345131
2434123132503213020552114345422130325432252500100321502342440435335025105043410
3131435231504421350241010405301043142100315114422021202242031254124153534242220
4433133144241014402311323204244421500313324504413224255000554124314003144021205
2531313333104553140035515413535204400124542304022511250024412143042402024524224
3022325455215430225105432551034135040520213000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000530430245232110055344
2520302

Counts Z_m of consecutive zeroes following position 3^m in the base-6 expansion of $\alpha_{2,3}$

m	3^m	Z_m
1	3	1
2	9	3
3	27	6
4	81	16
5	243	42
6	729	121
7	2187	356
8	6561	1058
9	19683	3166
10	59049	9487

Note that in the first 59049+9487 = 68,536 digits, there are at least 1+3+…+9487 = 14,256 zeroes (not counting zeroes in the "random" sections), so that the frequency of zeroes up to here is at least 0.208007… > 1/6). It can be shown that this ratio also holds in the limit. Thus $\alpha_{2,3}$ is not 6-normal.

A general nonnormality result for the Stoneham constants

Given co-prime integers $b > 1$ and $c > 1$, and integers p, q, $r > 0$, with neither b nor c dividing r, let $B = b^p c^q r$, and assume this condition holds:

$$D = c^{q/p} r^{1/p} / b^{c-1} < 1$$

Then the Stoneham constant

$$\alpha_{b,c} = \sum_{n=1}^{\infty} \frac{1}{c^n b^{c^n}}$$

is not B-normal. In particular, $\alpha_{b,c}$ is b-normal but not bc-normal.

Example: $\alpha_{2,3}$ is nonnormal base 6, 12, 24, 36, 48, 60, 72, 96, 120, 144, …

It is not known whether or not this result gives a complete catalog of the bases for which a Stoneham constant is nonnormal.

D.H. Bailey and J.M. Borwein, “Nonnormality of Stoneham constants,” *Ramanujan Journal*, http://crd-legacy.lbl.gov/~dhbailey/dhbpapers/nonnormality.pdf.

Nonnormality of the sum of two Stoneham constants

Assume that $\alpha_{b,c}$ and $\alpha_{d,e}$ are two Stoneham constants, both *B*-nonnormal as given in the previous result. Assume further that *c* and *e* are not multiplicatively related -- there are no integers *s* and *t* such that $c^s = e^t$.

Then the sum $\alpha_{b,c} + \alpha_{d,e}$ is also *B*-nonnormal.

D.H. Bailey and J.M. Borwein, "Nonnormality of Stoneham constants," *Ramanujan Journal*, http://crd-legacy.lbl.gov/~dhbailey/dhbpapers/nonnormality.pdf.

Open questions

- Is $\alpha_{2,3}$ 3-normal, or not? There are many open cases of this type.
- Is the nonnormality result given above a complete catalog of cases?
- $\alpha_{2,3}$ and $\alpha_{2,5}$ are each 2-normal. Is $\alpha_{2,3} + \alpha_{2,5}$ 2-normal? What conditions ensure that normal + normal is normal?
- Can the normality and nonnormality proofs of Stoneham constants be generalized to a larger class of real constants? Answer: Yes!
- Can absolute normality (i.e., b-normal for all integer bases $b > 1$) be established for a Stoneham-like constant? Or for any other constant?
- Can normality (to any base) be established for any of the "natural" irrational constants of mathematics – e, π, sqrt(2), log(2), zeta(3), etc.?
- Can nonnormality be proven for any of the natural irrational constants?

Any normality or nonnormality proof for a well-known mathematical constant would be a very significant result.

So would any result that helps to better understand why simple mathematical operations produce highly complex digit expansions.